AU/ACSC/0574/97-03

MODELING AND SIMULATION

IN SUPPORT OF TESTING AND EVALUATION

A Research Paper

Presented To

The Research Department

Air Command and Staff College

In Partial Fulfillment of the Graduation Requirements of ACSC

by

Major Tony Frank Hodge

March 1997

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (FROM - TO)
01-03-1997	Thesis	xx-xx-1997 to xx-xx-1997

4. TITLE AND SUBTITLE
Modeling and Simulation in Support of Testing and Evaluation
Unclassified

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Hodge, Tony F. ;

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME AND ADDRESS
Air Command and Staff College
Maxwell AFB, AL36112

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME AND ADDRESS
,

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APUBLIC RELEASE
,

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Testing and Evaluating are key to the successful development of an acquisition program. Operational Testing during the development phases of the acquisition process is the most difficult to accomplish. This is so because it requires that a system be nearly complete before you can conduct testing. Computer Modeling and Simulation of a proposed program is one way to achieve early operational test requirements. In this project, I will show how modeling and simulations support test and evaluation. At the conclusion of this paper, I will have given the reader the following: 1) an explanation of how similar programs can benefit from modeling and simulation, 2) an explanation of the current methods of validating models, 3) a discussion on why we should standardize validation methods, and 4) an explanation of how to standardize validation methods.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19. NAME OF RESPONSIBLE PERSON
			Public Release	38	Fenster, Lynn lfenster@dtic.mil
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER International Area Code Area Code Telephone Number 703767-9007 DSN 427-9007

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39.18

Disclaimer

The views expressed in this academic research paper are those of the author(s) and do not reflect the official policy or position of the US government or the Department of Defense.

Contents

Illustrations

Preface

This research project investigates how modeling and simulation support test and evaluation, throughout the acquisition process. This subject is of great interest to the acquisition community. It is of primary concern to testers who clearly understand the high costs associated with conducting field testing during development.

Having served in key positions on major Army acquisition programs, I have first hand knowledge of the high costs required to conduct field testing. Studies have shown that properly focused modeling and simulation can help to reduce program test costs without compromising operational requirements.

The overarching quest of this project will be to determine the feasibility of modeling and simulations. It will focus on how modeling and simulation (M&S) aid the material developer in identifying operational requirements during the program development cycle. Additionally, it will touch on other cost saving applications of M&S. This project was sponsored by the Defense Logistics Agency, as part of the Air Command and Staff College's student research program.

Abstract

Testing and Evaluating are key to the successful development of an acquisition program. Operational Testing during the development phases of the acquisition process is the most difficult to accomplish. This is so because it requires that a system be nearly complete before you can conduct testing. Computer Modeling and Simulation of a proposed program is one way to achieve early operational test requirements. In this project, I will show how modeling and simulations support test and evaluation. At the conclusion of this paper, I will have given the reader the following: 1) an explanation of how similar programs can benefit from modeling and simulation, 2) an explanation of the current methods of validating models, 3) a discussion on why we should standardize validation methods, and 4) an explanation of how to standardize validation methods.

Chapter 1

Introduction

The end of the Cold War has brought about many global changes. You could characterize it as a "New World Order." With this "New World Order" it is clear that we no longer require the defense structure that we needed in the past. However, it is also clear that we need a well-conceived security strategy, and a force structured to carry it out. There are two key factors affecting the United States (US) leader's ability to accomplish this mission: 1) the considerable reduction in defense funds, and 2) a new complex uncertain threat.

The reduction in funds and the new threat, has posed numerous challenges for US leaders. It has brought about a drastic downsizing and restructure in the military. It has meant less money for driving tanks, sailing ships, and flying airplanes. It has meant that operators and maintainers have had fewer opportunities to learn, maintain and improve their skills while they are on the job. It has meant fewer large-scale exercises. This has made it difficult for senior and mid-level leaders to gain the warfighting skills required to successfully use the new force structure effectively.

As leaders grapple with these problems, they have implemented many initiatives, to reduce defense costs, while attempting to maintain combat readiness. One of these initiatives is Acquisition Reform which provides a set of rules established by the DOD.

These rules basically say that acquisition personnel will attempt to get the best systems to the field, at the least possible cost.

Military Program Managers (PM) are chartered to implement acquisition reform initiatives within their programs. One of the major costs associated with developing a military system, is the cost to conduct operational field testing, prior to production of a system. As these costs continue to escalate, PM's must rely more and more on computer modeling and simulation, to assess system performance requirements during development.

This paper, will show how properly focused M&S can aid PM's in achieving their objectives. In setting the stage for this paper, I have included key terms and definitions. Chapters Two and Three will discuss operational requirements, and the acquisition process. An understanding of these two processes will better prepare the reader to understand how modeling and simulation aid the materiel developers in determining whether they are meeting operational requirements. Chapter four will provide an explanation of how similar programs can benefit from modeling and simulation. Finally, this project will discuss current methods of validating models and simulations, and how materiel developers could accomplish standardization of methods.

Definition Of Key Terms

Key terms relevant to this paper are modeling, simulation, validation, test, and evaluation. A model in its simplest state is a proxy. It is one entity used to represent some other entity for some well-defined purpose.[1] An example for a defense system would be a prototype. Models are developed and used to help hypothesize, define, or communicate some aspect of the original entity for which the model is a substitute.[2]

B. P. Zeigler in explaining his formal theory on modeling and simulation states that modeling is composed of five elements: the real system, the base model, the experiment frame, the lumped model, and the computer. The real system is the system being modeled. It is a source of observable data, in the form of input/output pairs. The base model is the investigator's image or mental model of the real system. It is the system that is capable of (at least hypothetically) accounting for the complete behavior of the real system. The experiment frame is the set of limited circumstances under which the real system is to be observed and understood for the purpose of the modeling exercise. It is a restricted subset of the observed output behaviors. The lumped model is a system that is capable of accounting for the output behavior of the real system, under the experiment frame of interest. It is a partial realization of the base model. The computer is the means by which the behavior of the lumped model is generated.[3]

Simulation is the development and use of computer models for the study of actual or postulated dynamic systems. An important aspect of simulation is the use of models for study and experimentation rather than the actual system modeled. The objective of simulation is to ensure that the computer faithfully reproduces the behavior implied by the model. Validation concerns the relationship between models and the real system. The objective of validation is to ensure that a model matches the system modeled. This is essential, because it's important to know that the conclusions about the model are reasonable conclusions about the real system as well.

Testing is a process used to obtain, verify, or provide data for determining the degree to which a system (or subsystem) meets, exceeds, or fails to meet its stated objectives. Evaluation is the review, analysis, and assessment of data obtained from testing. The purposes of test & evaluation are to: provide risk reduction or elimination, verify system parameters, and to support program milestone decisions. The two categories of test & evaluation are developmental and operational. Materiel developers conduct developmental test & evaluation to ensure that specified requirements are met and assist the design and development process. They conduct operational test & evaluation to measure operational effectiveness and suitability.[4]

Modeling and simulation, and test & evaluation both provide the materiel developer with necessary feedback to determine whether they are meeting the user's requirements. The focus of this paper will be to show how modeling and simulation can aid the materiel developer in conducting test & evaluation during program development. Thereby, eliminating some of the life cycle cost associated with field testing, while providing the user with a system that meets performance requirements.

Notes

[1]McGraw Hill, Encyclopedia of Science & Technology 11, MET-NIC, p. 312.
[2]Ibid., p. 312.
[3]Ibid., p. 312.
[4]Defense Systems Management College, *Fundamentals of Systems Acquisition Management*, 1994, p. 10-VG5-VG7.

Chapter 2

Operational Requirements

The requirements generation process is the first step in the acquisition process. It is the step in which users and combat developers formally identify warfighting deficiencies. It is important that they thoroughly conduct this process, because the operational requirements identified are the foundation upon which a new-start program is built. All efforts during development, especially T&E and M&S are geared towards achieving these operational requirements. Program Managers will rely on this feedback to make important program decisions.

The Department of Defense (DOD) uses two types of requirements documents, the Mission Need Statement (MNS) and the Operational Requirements Document (ORD). Users generate the requirements for defense systems, they are soldiers, sailors, airmen, marines, and civilians who will operate the system in the field. These users are found in commands, units, or elements that will use the system for mission accomplishment. The MNS describes warfighting deficiencies, and is the first document produced. It states requirements in broad operational terms, and is based on the user's analysis of assigned warfighting mission areas. The ORD is the second document produced. The user writes the ORD after they have studied alternatives to satisfy mission needs. They write the ORD to describe the proposed system.

Within the military, service level headquarters and subordinate commands assist and represent users in the development of the MNS and ORD. These representatives are unique to each service, and are commonly known as combat developers. Throughout the remainder of this paper, the term user will refer to these user representatives. For example, the Army's representative is its Training and Doctrine Command (TRADOC). TRADOC manages the Army's "combat developments" process for developing requirements.

During the requirements generation phase, the services combat developers conduct mission area assessments (MAA). They do this to identify mission needs of current and projected capabilities to defeat threats to national security. Mission areas are broad categories of warfighting responsibility, such as "fire support" for the Army, "strategic sealift/protection" for the Navy, "amphibious warfare" for the Marine Corps, and "air support and interdiction" for the Air Force.[1]

The decision to pursue a material solution is usually the last alternative available to fulfill a warfighting deficiency. Users normally attempt to fulfill their requirements through non-material solutions, before they pursue a material solution. Their initial step is to attempt to correct the deficiency through a change in training, doctrine, tactics, organization, or a combination there of. If they determine that this will not satisfy their requirement, then they must document it in a MNS, and start the process. It is standard practice to write a MNS for all mission needs that may result in an acquisition program.

The originator of the MNS will suggest its acquisition category (ACAT) determination and submit it for review, validation, and approval. The requirements validation authority will conduct this process. He will confirm or deny that a need exists and can only be resolved through a material solution. If he gives confirmation, he will give approval which means that the validation process is complete and the need is valid. The validation authority is also responsible for determining joint service potential. The final step is to send the MNS forward for a decision by the milestone decision authority (MDA).

Chiefs of the military services, heads of defense agencies, and commander-in-chief (CINCs) of unified commands, validate and approve their own MNS for potential non-major defense acquisition programs. The Joint Requirements Oversight Council (JROC), in the Joint Staff, is the validation and approval authority for potential major defense acquisition programs (ACAT I).[2]

The JROC designates the ORD validation and approval authority for ACAT I programs. Normally this is the Chief of a military service, the head of a defense agency, or the CINC of a unified command. ORD's for ACAT II, III, and IV programs are reviewed, validated, and approved by the Chiefs of the military services, directors of defense agencies and the CINCs. Service Chiefs can, and typically do delegate this authority to their operations deputy. Approved ORD's are submitted to the appropriate MDA for action. The ORD is used to develop the draft system specification for the Program Definition & Risk Management Phase.[3]

The developer updates the ORD at each milestone to incorporate the results of each acquisition phase. There is no need to update the MNS, because the ORD builds upon this initial document. The ORD should be thoroughly scrubbed at each milestone by the user and the development and test communities. They are checking to ensure that deficiencies and requirements are still valid when compared to the latest threat, defense planning guidance, and military strategy documents.

Notes

[1]Defense Systems Management College, *Fundamentals of Systems Acquisition Management*, 1994, p.11-2.

[2]Ibid., p.11-3.

[3]Ibid., p.11-3.

Chapter 3

Program Development Process

Materiel developers need to completely understand the life cycle process in order to effectively integrate M&S into their T&E programs. An understanding of the process will allow early integration of M&S and thereby maximize your benefits. According to some experts, developers can benefit from M&S as early as phase 0 of the acquisition process. Early M&S allows analysts to explore different concepts, without wasting time and money building hardware. It also gives them information to identify and eliminate unrealistic solutions, without expending unnecessary time and money.[1]

The DOD acquisition Lifecycle process has four phases and milestones points. The four phases are: concept exploration & definition (CE&D), program definition & risk reduction (PD&RR), engineering & manufacturing development (EMD), and production, fielding, deployment & operational support (PFD&OS). The four milestone (MS) decision points are: MS 0, MS I, MS II, and MS III.

The first phase of the acquisition life cycle process, is CE&D. This phase begins immediately following a successful MS 0 review. Prior to MS 0 is the requirements generation phase. MS 0 is the first MS decision point. At this point, the milestone decision authority (MDA) will authorize the study of a minimum number of alternative solutions, hence starting the CE&D phase.

The primary objective of this phase is to identify system concepts to satisfy the mission need. The focus is on competitive exploration of potential solutions, while working with industry and the user to determine trade-off in cost, schedule and performance. The developer will also look at current ongoing development efforts, and non-developmental items to satisfy user requirements. The user will refine his Operational Requirements Document (ORD) to describe the selected system concept prior to MS 1.The ORD is the basis for the draft system specification for the next phase.

Two key activities that will occur in this phase are: initiating and updating program documents, and establishing program exit criteria for the PD&R phase. The program exit criteria that must be established, should be major “show stoppers” that require intensive management to ensure the program is ready to proceed past the next MS.

Once these and other key activities are conducted, the results are presented to the MDA at a MS I review. If the MDA approves, he issues the acquisition decision memorandum (ADM) authorizing the initiation of a new defense acquisition program, and procedure into the next phase.

The next phase of development is the PD&R phase. During this phase, contractors initiate design on alternatives selected from the previous phase. The contractors will select the systems best capable of fulfilling the mission need. The program manager may preserve competition, through a Competitive Prototyping Strategy, to offset some of the risk associated with the decision. The primary objective of this phase is to select a system for development and production. Other objectives of this phase are to test processes and technologies to identify risk areas, and to develop risk abatement plans. The contractors will consider and integrate supportability and manufacturing design issues as they start system design.

The program manager (PM) has several key activities that he must ensure is accomplished. One primary activity, is the fabrication of prototype systems or subsystems to support design, development, and test and evaluation to identify areas of risk. Contractors will analyze developmental test results to determine the degree of risk remaining to new or emerging technologies. The PM must identify, through prototyping and testing, trade-off required to maximize cost, schedule and performance benefits to the government. Competition may be limited to design competition, subsystem/component competition, or full system level competition. ACAT I programs must use a competitive prototyping strategy, or obtain a waiver from the USD(A&T).[2] This is an opportune time to insert M&S technology. It can be used to assist developers in testing processes and technologies to identify risks, without building prototypes.

The PM ensures that previous activities are completed, to help select a system for EMD. He will ensure that documentation to support the MS II decision is prepared summarizing the results of the PR&D phase. Finally, the MDA will hold a MS II decision review. A favorable decision at MS II authorizes the program to proceed into EMD, and approves the Development Baseline and exit criteria for EMD.

The next phase of development, is EMD. The primary objective of this phase is to complete system development. This is a difficult task because typically, all parts of a system will not have been developed. The materiel developers must firm up the initial system design through a series of development and operational tests, to ensure that they are meeting the user's operational requirements. This is another opportunity to benefit from M&S technology.

A key activity that takes place in this phase is the development and procurement of production representative systems to support test and evaluation and to evaluate the contractor's ability to produce the end item. Low Rate Initial Production (LRIP) may be conducted to help reduce technical risk. LRIP, provides items for operational testing, establishes an initial production base, and allows for an orderly ramp-up to full rate production. The goal of LRIP is to produce and test systems that provide a realistic portrait of performance under operational conditions.[3]

The end products of the EMD phase will be a design matured and ready for full rate production, and an updated acquisition program baseline for proposed cost, schedule and performance objectives and thresholds. The production phase will begin immediately following a favorable MS III decision, and production contract award.

The next phase of development is the PFD&OS phase. The program managers primary objective during this phase is to ensure that systems are produced at an economical rate and deployed in accordance with the user's requirement. The developer may conduct follow-on operational test and evaluation (FOT&E) to confirm that production units meet operational effectiveness and suitability requirements. The Operations and Support Phase starts as soon as the first system is fielded. It continues until the system leaves the inventory.

Key activities in this phase are, manufacturing, contract monitoring, and acceptance testing. The program manager will rely heavy on Defense Contract Management

Command (DCMC) personnel for contractor surveillance. Also, user trainers will undergo extensive training to learn how to operate and maintain the system.

Other production activities include production acceptance test and evaluation, monitoring of the customer's quality assurance program, and adhering to a production schedule that meets the system's Initial Operational Capability (IOC) and Full Operational Capability (FOC).

Key activities that will take place during the Operations and Support portion of this phase are update of the configuration baseline and the system threat assessment. The user's will monitor fielded systems to assess the effects of aging on system capabilities. When appropriate, modifications will be undertaken to enhance operational performance, reduce costs, and improve supportability.

Finally, there is MS IV, a MS decision held for selected programs to determine if major modifications are necessary for systems still in production. At this point, the acquisition Lifecycle potentially starts over, if determined by the appropriate authority. Figure 1 depicts the acquisition life cycle process.

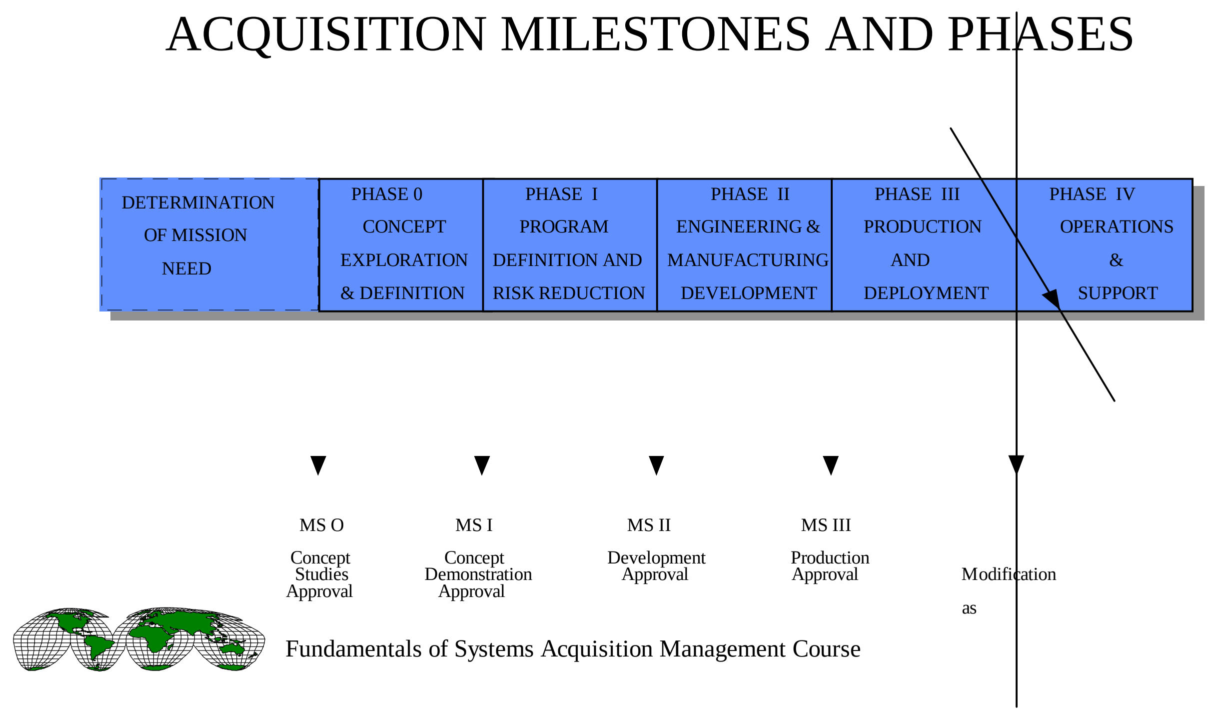

Figure1. Acquisition Lifecycle Process

Notes

[1]Neal T. Lovell, Computer Sciences Corporation, interviewed by author, Advisory Staff Operation Research Technology Development Center, Huntsville AL, 3 January 1997.

[2]Defense Systems Management College, *Fundamentals of Systems Acquisition Management*, *1994*, p.13-2.

[3]Ibid., p.14-3.

Chapter 4

Similar Programs Benefits Form M&S

Materiel developers have historically conducted testing of new systems using a test-fix-test approach. They usually begin at the component and subsystem levels and increase in complexity up to and including system level field testing. They use modeling and simulation extensively to investigate and solve complex technical issues. This approach worked well in the past. However, with the increased complexity and high cost of new weapon systems, a change must occur. Developers must change their approach to test & evaluation, modeling & simulation, as well as procurement of systems. Test & evaluation and modeling & simulation efforts should be integrated and implemented early and throughout the entire system Lifecycle. It should be conducted in a systematic, holistic, and consistent manner.[1]

One agency that has clearly received this message is the US Army's Test and Evaluation Command (TECOM) Redstone Technical Test Center (RTTC), Redstone Arsenal AL. RTTC initiated extensive efforts to integrate M&S with T&E in 1992. They began their efforts, after receiving a challenge from then Deputy Under Secretary of the Army Stephen Convers. The challenge was to "integrate M&S into every applicable area," of development.[2]

RTTCs response to this challenge was envisioned and spearheaded by Mr. Larry Johnson, the Director of RTTC. Mr. Johnson envisioned an integrated Test/M&S process that would integrate the efforts of system designers, modelers, testers, and evaluators to produce an efficient thorough integrated test plan. This approach will reduce the time, resources, and risks of the acquisition process. The capstones of this vision are:

- Develop an integrated modeling, simulation and test plan which supports the design and evaluation process from beginning to end.
- Utilize M&S to evaluate component and subsystem tests at the system level with and without external disturbances (operating environments).
- Utilize component, subsystem, and system level testing to validate models and simulations.
- Use M&S to enhance, but not eliminate testing [3]

RTTCs tool for executing this concept is their Virtual Proving Ground (VPG). The VPG is a cohesive and comprehensive capability for testing concepts, virtual prototypes, hardware prototypes, subsystems, and full systems. It contains standardized automated test methodology, synthetic stimuli and environments based on TECOM Ground Truth data and physics. The VPG is a distributed synthetic test environment that may be located and used wherever it makes sense. RTTC envisions the VPG providing the following: 1) an integrated, inter-connected testing resource with rapid data acquisition, reduction and reporting, 2) interconnection between the TECOM Test Centers or distributed testing; and 3) remote access to both live and virtual test facilities. Figure 2 depicts the VPG vision. RTTC is developing the VPG under a cross-Test Center Integrated Product Team format. Each TECOM test center has lead responsibilities for the various functional working areas. The functional working areas are: interface architecture, databases, synthetic stimuli, automated test procedures, virtual instrumentation, verification, validation and accreditation (VV&A), computers and communications. The commodity working areas are: ground systems, aircraft, small missiles, large missiles, chemical/biological, and command control/ computers/ communications and intelligence (C4I).[4]

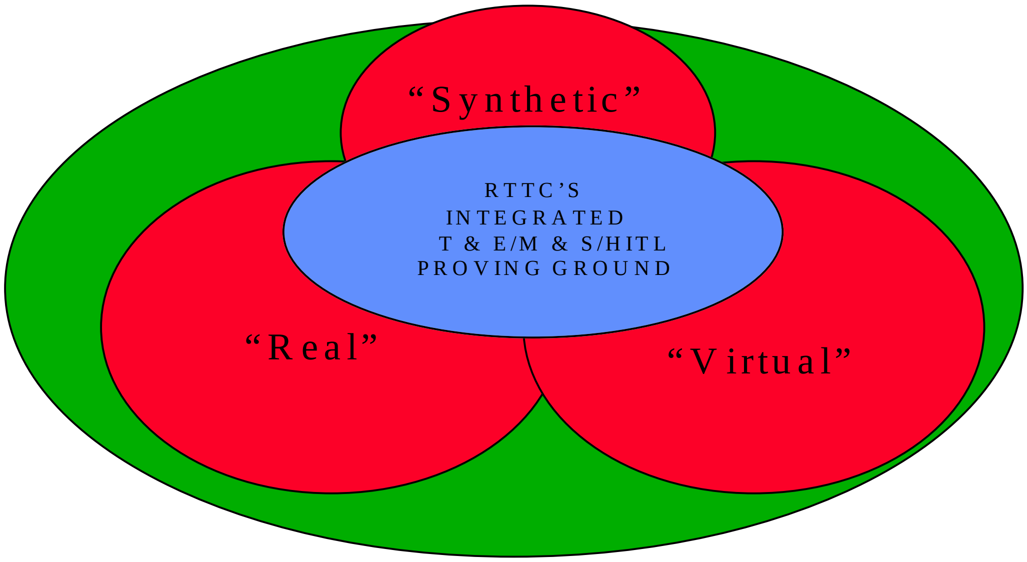

Figure 2. Proving Ground

RTTC demonstrated its VPG concept on three of its missile programs. They used the Javelin missile program as a pathfinder to demonstrate how the VPG could be applied to a program nearing production. They used the Follow-on-to-Tow missile program to demonstrate how the concept could be applied to a new start program. Finally, they used the Longbow Hellfire program to highlight the benefits of M&S on a program in Low-rate Production.

RTTC applied M&S to the Javelin programs Stockpile Reliability Program (SRP) as the program was nearing the production phase. They added requirements to the SRP plan to "conduct a simulation program in conjunction with the various testing programs to provide a better analysis of the overall missile performance." They used six degree of freedom simulation to predict missile performance in various scenarios to include obscurants and component degradation. RTTC performed a hardware-in-the-loop (HWIL) simulation on the items that were not one shot. This application of VPG uses standard test data injected into the all-digital weapon system simulation to project system level performance. It uses component, subsystem, and system level HWIL testing to evaluate performance. It includes tactical environments and their effects, and provides for connectivity to the thermal sighting unit as well as the virtual prototype and trainer.[5] This application of the MVPG clearly demonstrates how developers can integrate M&S into the development process. This application saved money through non-destructive testing and fewer destructive flight tests.

RTTCs approach to applying its concept to the FOTT missile program, is to implement Simulation Test & Evaluation Plan (STEP) in the program from beginning to end. They plan to do this by integrating simulations, testing, and design of experiments in order to optimize the risk/cost relationship. Their objectives are: 1) define FOTT data sources, 2) identify alternatives to "the norm" to integrate the concepts of STEP, 3) explore possible cost savings that may be achieved through the STEP implementation, 4) integrate STEP into the FOTT Engineering Manufacturing and Development Contract Statement of Work, and 5) to integrate STEP paths in the system evaluation plan to support program milestone decisions.[6] The bottom line is that the FOTT IPT's mission is to integrate the simulation plan, the master test plan, and the evaluation plan through a design of experiments. The final product will be a single Integrated Simulation, Test and Evaluation Master Plan. The benefits of this approach are unlimited. According to experts, inserting M&S early into a program, will yield the developer the following benefits: 1) the ability to explore numerous concepts without building hardware, 2) the ability to economically refine designs and identify workable solutions, 3) the ability to test the system under various conditions, and 4) the ability to repeat testing at a much lower cost than conducting live testing.[7]

RTTCs application of M&S to the Longbow Hellfire program resulted in a yearly cost avoidance of greater than $5 million dollars. RTTC developed an alternative method to conduct lot acceptance testing of Longbow Hellfire missiles. Previously, they would conduct fly-to-buy testing. This involved taking a random lot sample and flight testing to determine lot acceptance. The cost per round was extremely high. RTTCs solution to this problem is their Simulation/Test Acceptance Facility (STAF). Utilizing the STAF, RTTC uses a combination of open and closed-loop testing to fully characterize the All Up Round (AUR) missile under test. The open-loop testing involves characterizing the control actuator system, the inertial measurement system and the end-to-end RF chain. The closed-loop testing involves presenting the missile under test with in-band threat and background scenery, real-time three-axis motion in pitch, yaw, and roll and injected inertial measurement data. This process occurs until simulated target impact. RTTCs solution allows real-time flight dynamics, real-time threat and background scene generation, and comprehensive data collection to the point of simulated target impact. They perform all tests under extreme temperature conditions to simulate various climates.[8]

As stated earlier, this method has resulted in great cost savings. This savings has resulted primarily because of the following reasons: 1) fewer flight tests, augmented by HWIL simulation testing at the AUR level, 2) nondestructive testing, allowing the AUR to be placed in inventory after testing, and 3) testers can conduct multiple test scenarios at various launch conditions, ranges, targets, and temperatures.

RTTCs proposed concept to integrating M&S with T&E is right on target. Their tool for execution, the VPG is an idea whose time has come. Their integrated modeling and hardware testing will support the total weapon system Lifecycle. They must develop and validate models through testing during the research and development phases. If they model and test in this manner, they can develop model libraries and databases. Materiel developers can use these libraries and databases to model and study concepts before actually building hardware prototypes, for all types of systems. As shown, these efforts combined, can yield excellent results.

Notes

[1]Charles M. Crocker and Larry M. Johnson, "Cost Effective Weapon System Development Through Integrated Modeling and Hardware Testing" (Redstone Technical Test Center, December 1993), p. 1.

[2]Charles M. Crocker, "Application of the Simulation, Test, and Evaluation process to a Tactical Weapon System" (Redstone Technical Test Center, December 1996), p. 1.

[3]Ibid., p. 1.

[4]Ibid., p. 2.

[5]Ibid., p. 5

[6]Ibid., p. 6.

[7]Neal T. Lovell, Computer Sciences Corporation, interviewed by author, Advisory Staff Operation Research Technology Development Center, Huntsville AL, 3 January 1997.

[8]"Use of Modeling and Simulation to Reduce Missile Acquisition Cost," *RD&A*, November-December 1996, p.23.

Chapter 5

Validation Methods

There are many challenges facing developers in validating military simulations and models. This subject is very difficult and complex, and has many different levels and interpretations of validity. For example, a model has face validity simply if it is accepted as reasonable for its intended purpose by people who are knowledgeable about the system under study. A model is replicatively valid if its trajectories match the real system input/output data used in its development. A model is predictively valid if its trajectories also match experimental data not used in its development. A model is structurally valid if it is behaviorally valid also matches the structure of the base model.[1] In other words, if it generates behavior within its experiment frame in substantially the same way that the real system is believed to generate this behavior.

In this chapter, I will discuss validation methods currently being used. I will compare and contrast these methods, and show how they can be integrated to make one standardized method. The methods that I've chosen to discuss cover the full breadth of validation. They are works by Osman Balci, and Averil Law and David Kelton.

Balci's method implies that validation is a process that should be implemented throughout the entire life of a simulation. His methodology is based upon the following six principles:

- Validation is not a 'yes or no' question.
- Model validation should be conducted throughout the Lifecycle of the model.
- Validation requires independent analysis to prevent any biases of the model developer.
- Validation requires creativity and insight into the problem facing the analyst.
- Complete testing of a model is not possible.
- Validation must be planned and documented.[2]

Balci asserts that validation and simulation is an iterative process. Also that it is expected that the analyst may have to revisit a previous step should an error be discovered. His methodology consists of the following nine steps:

1. Problem formulation
2. System and objective definition
3. Model qualification
4. Communicative model
5. Programmed model
6. Experiment design
7. Data validation
8. Model validation
9. Presentation validation

Problem formulation is the first phase of his method. In this phase, the analyst goes through the process of determining if the problem formulation is identical to the actual problem. If the formulated problem does not contain the actual problem, the analyst has committed a type III error, solving the wrong problem.[3] At this point, the analyst will investigate solution techniques, and choose the appropriate technique to solve the problem. If he chooses simulation, he will continue along the Lifecycle process.

The next step is system and objective definition. This step is used to determine system characteristics. The analyst validates six major system characteristics that tend to cause failures: 1) system changes, 2) system environment, 3) counterintuitive behavior, 4) drift to low performance, 5) interdependencies, and 6) interdependencies. The third step is model qualification. It is the process of justifying the appropriateness of the conceptual model. Balci defines the conceptual model as the model formulated in the mind of the analyst. Model qualification can also be described as the process of justifying the assumptions that the analyst has postulated for the model.

The next step, the communicative model is the model representation that can be communicated to other people. It can be judged against the real-world system, the study objectives, and the study constraints. The following step, is appropriately called programmed model verification. It is the computer executable code.

The sixth step is experimental design. This is the process of creating the experiments, or scenarios, with which the model will be exercised. Validating the experimental design is to evaluate the appropriateness of the scenarios for use to achieve the goals of the simulation analysis. The seventh step, data validation is the process of checking that the input data is accurate, complete, unbiased, and used in the proper context for the model.

The eighth step, model validation, is the process of checking that the experimental model is appropriately accurate to fulfill the study's objectives. The experimental model is the programmed conceptual model coupled with the designed experiments and the valid data. The final step is presentation validation. This is the process of justifying that the output results are interpreted, documented and communicated with appropriate accuracy. Documentation is an extremely important factor in presentation validation.[4] Balci's method is an iterative process. The analyst performs the aforementioned steps until the objectives are met, or until the objectives are determined unattainable.

Law and Kelton's method is a three step process. The three steps are:1) develop a model with high face validity, ensure that it looks reasonable to system experts, 2) empirically test the model assumptions, and 3) is to determine how closely the model output data resembles the expected (real-world) output data.[5]

Law and Kelton suggest the following steps to achieve high face validity:

- Conduct extensive interviews with system experts.
- Collect data from a system similar to the one being modeled.
- Use established, relevant theories.
- Use relevant results from similar simulation models.
- Use experience and intuition.
- Maintain continuous dialogue with the customer/client throughout the study.
- Perform a walk-through of the conceptual model to all key people.

The second step suggests empirically testing the model assumptions. To accomplish this, they suggest: 1) testing the probability distributions used, and 2) sensitivity analysis on output data.

The third step is to determine how closely the model output data resembles the expected (real-world) output data. Law and Kelton suggests that when model outputs do not agree exactly with real system output data, calibration may be a factor. Calibration may be added or multiplied to possibly correct absolute output. They warn that analysts should use caution when using a calibration factor. Although calibration may achieve proper results for one set of inputs, the model might not be valid over the entire range of inputs. Law and Kelton's approach to validation is a combination of common sense, subjective tests, and empirical tests. They stress that empirical tests of output data are the most definitive tests for validation.

Balci's validation methodology is far more comprehensive and detailed than that of Law and Kelton's. Balci separates the validation effort into eight specific types of validation defined at different times of the Lifecycle. He recommends many techniques to use to achieve positive validation in the particular validation classifications. Law and

Kelton present a more general methodology. Both methodologies are similar in that they both stress empirical testing as the primary means of achieving validation. Both works include subjective techniques, primarily face validation, as an important, but secondary tool. Both stress the necessity of performing the validation process throughout the entire Lifecycle of the model, and detailed documentation of the process.

Notes

[1]McGraw Hill, Encyclopedia of Science & Technology 11, MET-NIC, p. 312.

[2]Michael R. Elmer, "Issues and Challenges in Validating Military Simulation Models" (research paper Air University, December 1995), p. 7.

[3]Osman Balci, *Validation, Verification, and Testing Techniques throughout the Life Cycle of a Simulation Study,* (Virginia Polytechnic Institute, 1994), p. 97.

[4]Michael R. Elmer, "Issues and Challenges in Validating Military Simulation Models" (research paper Air University, December 1995), p. 10.

[5]McGraw Hill, Simulation Modeling and Analysis, 1991, p. 62.

Chapter 6

Validation Standardization

To state the bottom line up front, the DOD should standardize its M&S validation methods. This seems to be a difficult task, because of the complexity of validation, and the differing views on how to properly conduct validation.

Standardization of DOD validation methods would provide the DOD with several benefits. Some of these benefits are: 1) a standardized triservice validation method, 2) a standard data validation data base, and 3) savings in time and money.

The first, and what I consider most important benefit is a standardized triservice system. Interoperability has become an essential consideration in military weapons development. Establishing a triservice validation method would greatly contribute to achieving interoperability. The Defense Modeling and Simulation Office (DMSO), was primarily established to promote the use of Interoperable standards and protocols in modeling and simulation. The DMSO was also established to increase DOD cognizance of, and facilitate coordination among the departments M&S activities; and to stimulate joint use, high return M&S investment.[1] A standardized triservice validation method would be in keeping with the goals of the DMSO, especially in achieving interoperability.

The two other benefits listed are both residuals from establishing a triservice method. A natural out-flow from a standardized system, would be a standardized data base. This would also contribute to assisting the DMSO in achieving its goal of interoperability, by creating a centralized, standardized data base.

The third benefit is a savings in time and money. A standardized process would ensure consistency among the services, and maximize efficiency of the validation process. This would thereby create great cost savings.

Materiel developers can achieve standardization of the validation process. There are certain characteristics that are common to all processes. They must simply identify these characteristics, and integrate and implement them into a organized process. In conducting the research, to support this project, seven steps appeared consistent through most methods, and therefore could answer the mail for a standardized process. These steps are steps proposed by Robert Sargent, who contends that as a minimum the following steps should be apart of any validation process:

- The analyst and customer should agree before the study begins on the basic validation approach.
- The assumptions and underlying theories of the model should be tested.
- Face validation of the model should be checked on the conceptual model with each model iteration.
- The model's behavior should be checked with the computerized model on each iteration.
- The analyst should compare the model and system behavior for at least two sets of experimental conditions.
- The analyst should fully document the validation process.
- Schedule periodic reviews of the validation, if the model will be used over time.[2]

I completely agree with Sargent in that at a minimum, these steps should be present in any validation process. The analyst can of course develop a much more detailed list if necessary. The level of technical difficulty will determine the level of validation. In summary, I feel that the DOD should standardize on validation. They can do this by establishing DOD policy that governs its implementation and execution. Currently, the

DMSO is promoting interoperability standards and protocols in M&S. I feel that M&S validation fall under their scope of work. If the DMSO create a standardized triservice validation method, it will result in the previous mentioned benefits, and further the other goals of the DMSO.

Notes

[1]Charles M. Crocker and Larry M. Johnson, "Cost Effective Weapon System Development Through Integrated Modeling and Hardware Testing" (Redstone Technical Test Center, December 1993, p. 1.

[2]Minutes of the Some Comments on Model Validation, Computer Modeling and Simulation Conference conducted at Simulations Councils Inc., La Jolla CA, 1982, p.36.

Chapter 7

Conclusion

Protecting US interests in the post Cold-War world will continue to produce many challenges. The DOD is facing these challenges head-on with diligent, meticulous planning. The DOD's goal is to maintain a smaller military force that is the best equipped and trained in the world. Science and Technology will play a key role in achieving this goal, especially M&S technologies. The focus of this paper has been M&S technology, and its application in weapons system development.

In setting the stage for this paper, I gave a detailed explanation of the requirements generation process, and the program development process. An understanding of both is key to initiating and effectively managing an acquisition program. This is true because operational requirements are the foundation upon which programs are built. They define the warfighting deficiencies of the users. Understanding the development process is essential because material developers need to know the process in order to plan for and integrate M&S throughout the entire process.

As I have proven in this paper, there are many benefits to integrating M&S into T&E. In chapter four, I explained how similar programs can benefit from M&S. I discussed three programs, showing their similarities and explained how to implement M&S in programs at different stages of the Lifecycle.

The DOD's lead agency in implementing M&S / T&E, is the US Army's TECOM RTTC. RTTC has established an initiative to use component and subsystem level test data to assess system level performance. Their objective is to develop quality material at lower cost, meeting program milestones with fewer hardware tests. They also plan to use test data to facilitate the design process by providing the evaluation of system performance and system reliability early in the design process. A method has been proposed to programmatically implement this initiative by early T&E and M&S integration in the planning and scheduling phase, and by specifying the use of standard M&S tools that will allow for the insertion of component level test data.[1] In chapters five and six, I discussed M&S validation. I discussed two of the current methods of validation, comparing and contrasting them. Finally, I discussed how the DOD could standardize on validation. I concluded that yes, we can and should standardize on validation.

The bottom line is yes, M&S does support T&E. Material developers can maximize the benefits of M&S, if they plan and integrate it up front and throughout a programs Lifecycle. RTTC has clearly demonstrated that it can be done. The challenge to the rest of the DOD acquisition community, is to "just do it."

Notes

[1]Charles M. Crocker and Larry M. Johnson, "Cost Effective Weapon System Development Through Integrated Modeling and Hardware Testing" (Redstone Technical Test Center, December 1993, p. 2.

Bibliography

Balci, Osman Validation, Verification, and Testing Techniques throughout the Lifecycle of a Simulation Study. Virginia Polytechnic Institute, 1994.

Cochrane, CB Fundamentals of Systems Acquisition Management (FSAM) coursebook, Defense Systems Management College, January 1994.

Crocker, Charles M. "Application of the Simulation, Test, and Evaluation Process To A Tactical Weapon System." Unpublished report, US Army TECOM, RTTC, Redstone Arsenal AL, December 1996.

Crocker, Charles M. and Johnson, Larry H. "Cost Effective Weapon System Development Through Integrated Modeling and Hardware Testing." Unpublished report, US Army Test and Evaluation Command (TECOM), Redstone Technical Test Center (RTTC), Redstone Arsenal AL, December 1993.

Elmer, Michael R. "Issues and Challenges in validating military simulation models." Unpublished research paper Wright-Patterson AFB, Ohio December 1995.

Johnson, James B. and Ray Jerry A. "Use of Modeling and Simulation To Reduce Missile Acquisition Cost." ARMY RD&A 23-26, November-December 1996.

Law, Averil and Kelton David. *Simulation Modeling and Analysis.* McGraw-Hill, New York, NY, 1991..

Lovell, Neal T. Advisory Staff Operation Research Technology Development Center, Computer Sciences Corporation, Huntsville AL: Personal Interview, 3 January 1997.

McGraw-Hill Encyclopedia of Science & Technology, 11 & 17 MET-NIC & SAB-SON.

Sargent, Robert. "Some comments on Model Validation Computer Modeling and Simulation." Minutes of simulation conference,. Simulation Councils Inc., La Jolla, Ca, 1982.